



生命機能物理学講座 特任助教

鮑 雨

ホウ ウ
博士 (情報学)
Project Assistant Professor
Division of Physics for Life Functions

Bao Yu

Ph.D. in Informatics

バイオインフォマティクス/
機械学習/
深層学習/最適化

Bioinformatics/Machine learning/
Deep learning/Optimization

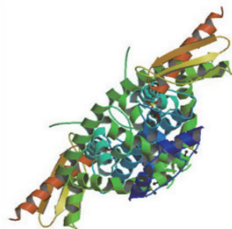
研究概要

近年、バイオサイエンス分野と薬学分野において利用されるデータサイエンスの技術に著しい革新が起っています。それに伴い、薬学研究等におけるオミックスデータから、がん診療データまでの幅広い分野において、膨大なデータが蓄積されてきています。

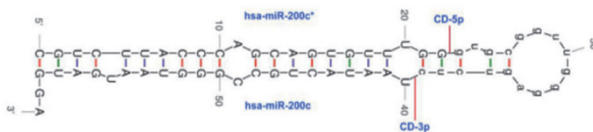
しかしながら、これらの分野のビッグデータ活用には、様々な工夫や手間が必要になります。そのため、機械学習を用いて収集した膨大なデータをクラスタリングし、有用な情報を抽出する技術の開発、そして、その有効性の検証を行うことに期待が集まってきています。

深層学習、人工知能 (AI) は多層パーセプトロンという階層型ニューラルネットから発展し、近年、画像認識や自然言語を対象とした問題に対し、他の手法を圧倒する高い性能を示すことが分かってきており、また、バイオインフォマティクスの分野における様々な応用例も報告されてきています。

私は、人工知能 (AI) における機械学習や深層学習のアルゴリズムを、化学・生物・薬学のデータに適した形で構築する研究開発を行っており、その中でも特に、タンパク質やRNAの切断部位を予測するアルゴリズムの開発や生物情報ネットワークを対象にした、有用な情報を抽出するための、深層学習のアルゴリズムの研究開発を中心に進めています。また、現在、慶應義塾大学では、中分子創薬のライブラリー開発に役に立つAIの開発を目指しています。



3D structure of Dicer enzyme



Calculated secondary structure of MicroRNA

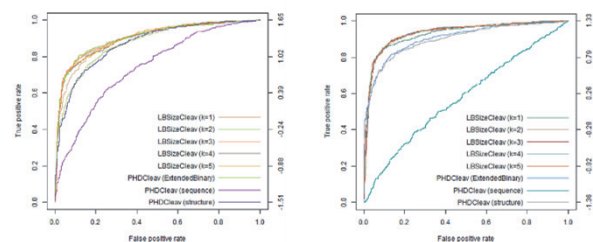
In the past twenty years revolutionary changes have been observed in data science regarding bio science and pharmacy areas. A rapid explosive of data has been published ranging from data generated in pharmaceutical studies to data generated in cancer analysis. However, how to deal with these data correctly and efficiently to achieve meaningful results still remains a challenging problem.

Artificial Intelligence has gained an increasing attention in the development of Bioinformatics and data science, partially benefits from the fact that a large amount of data becomes available recently. In these AI technologies, Deep learning has emerged and proved its value in dealing with data with an extremely large size in computer vision and speech recognition, my research focuses on the development of machine learning/ deep learning algorithms regarding various biological/pharmacology problems including prediction of cleavage site specificity for both proteins and RNAs/ analyzing various types of biological network and extracting useful information.

主な論文

Bao, Y., Hayashida, M., Akutsu, T. 2016. Lbsizecleav: improved support vector machine (svm)-based prediction of dicer cleavage sites using loop/bulge length. BMC bioinformatics. 17(1), 487.

Bao, Y., Marini, S., Tamura, T., Kamada, M., Maegawa, S., Hosokawa, H., Song, J. and Akutsu, T., 2018. Toward more accurate prediction of caspase cleavage sites: a comprehensive review of current methods, tools and features. Brief Bioinform, p.bby041.



ROC curves for LBSizeCleave compared to the PHDCleave